Journal of Information Systems Engineering and Management
2025, 10(2)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Louvain Modularity and Ensemble Gradient Boost Absolute Congruence Deep Belief Classifier for Legal Documents

Nineesha P[1], Dr. P Deepalekshmi[2]

[1] *Assistant Professor, Department of computer science and Engineering, Sree Buddha college of Engineering, Pattoor, Alapuzha, Kerala. p.nineesha@gmail.com*

[2] *Professor, Department of computer science and Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Tamilnadu. deepa.kumar@klu.ac.in*

ARTICLE INFO

Received: 07 Oct 2024

Revised: 12 Dec 2024

Accepted: 20 Dec 2024

ABSTRACT

As far as the Indian law system is concerned, judgment is contemplated as a final decision that is stated by the court for a stipulated case. Over the past few years, owing to the mushrooming growth in technologies, like, machine learning, deep learning and ensemble learning, judgments are kept in digital forms for ease of use. With the emergence of an automated system, the proceedings are made simple for the law professional. Several law firms have already started applying these learning mechanisms that specifically concentrate on emphasizing information from the Indian judgmental document. Also extracting key factors from legal documents is considered as a time consuming process. To ease the task of classifying legal documents, a method called Louvain Modularity and Ensemble Gradient Absolute Congruence (LM-EGAC) deep belief classifier for legal documents is proposed. The LM-EGAC method is split into two sections, namely optimized topic modeling (i.e., extracting optimized topic words) and classification (i.e., identifying case type). First, Highly Associative Unique Keyword and Louvain Modularity-based Optimized Topic Modeling are applied for extracting optimal topic words for further processing. Followed by which the Ensemble Gradient Absolute Congruence boosting model is applied for classifying the legal documents into corresponding cases. The Ensemble Gradient boosting algorithm initially forms an Absolute Congruence function with the purpose of classifying input legal documents along with the optimized topic words. Next, the results of the deep belief classifier are joined to construct a strong classifying by minimizing the error with the aid of binary entropy function and Louvain Modular function. The proposed LM-EGAC method is experimentally validated with several performance factors like, precision, recall, error and computational time. The experimental results quantitatively confirm that the proposed LM-EGAC method achieves better precision with minimum error and computation time upon comparison with the traditional methods.

Keywords: Louvain Modularity, associative Unique Keyword, topic Modeling, Ensemble Gradient Boost, Absolute Congruence, Deep Belief Classifier.

1. Introduction

For the last few decades, evolution of computational intelligence in Indian judgmental documents has been in the process with the inception of sophisticated techniques. Also elaborate analysis on extensive types of subject matters like, case prediction and case classification is being executed on Indian judgmental documents. Design of an automated constitution can bring forth assistance in law negotiation for legal professionals and also aids as a helping device for the general public on comprehending his/her case.

One of the power mechanisms for acquiring hidden patterns in exhaustive documents is modeling the concerned topic. With the employment of topic modeling, themes can be recognized that are found to be highly associated, therefore resulting in certain regions while judging for both spatial and temporal convolution. A Latent Dirichlet Allocation-based natural language processing was proposed in [1] for obtaining unique critical topics concerning different types of legal cases. By applying these mechanisms, the method of controversial legal issues can be solved with reasonable accuracy. Despite improvement observed with reasonable accuracy, nevertheless, lack of optimized topic models for legal data results compromise the information being retrieved or the overall precision and recall in a significant manner.

Artificial Intelligence (AI) procedures like Machine Learning (ML) have not been explored to their extreme possibilities in the legal area. This has been moderately owing to the inadequate clarifications they issue regarding their decisions. However, computerized expert mechanisms with evaluative potentialities can be particularly helpful when legal practitioners explore charters to accumulate circumstantial knowledge for their cases. Hence, a hybrid method that applies Multi-label Classification using ML was proposed in [2]. Here, both the Natural Language Processing mechanisms and deep legal reasoning were explored with the purpose of acquiring the entities, like the parties taking part in. This type of hybrid method resulted in the improvement in terms of precision, recall and accuracy considerably. Despite improvement observed with respect to precision, recall and accuracy, the computational or training time involved was not focused. To address on this accuracy and training time aspects, in our work, we intend to use Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier.

With the emergence of the epoch of artificial intelligence, several materials and methods constituted by machine learning and deep learning have exquisitely replaced general public lives and work. Also driven by enormous data size and artificial intelligence, Indian judicial institutions at all extent entered legal AI with the provisioning of smart judicial services. Moreover, the evolution of AI to the judicial domain is hitherto an inclination and leads the way to new progress and changes to judicial services. However, acquiring the most pertinent articles is yet considered as a major task.

In [3], a multi-label classification method for classifying judicial documents was presented. Here, by means of multi-label classification ensured improvement in terms of classification accuracy. Yet another deep learning technique was applied in [4] for patent classification owing to its laborious and time consuming process. By applying deep learning the coverage error was said to be reduced with improved precision.

Analysis of legal documents in an automatic fashion nowadays is considered to be mandatory for ensuring smooth and robust legal practice and research. As a matter of course, analysis of data should be performed in such a manner so that a prospect and opaque mechanism is followed thoroughly. These necessitates are additional significant with respect to legal data. This reinforces that lawyers, judges and legal decision-makers engage with and the cost of error paves the way for identifying an automatic processing and analysis.

On the basis of an enormous amount of judicial cases, a pre-training language model Bidirectional Encoder Representations from Transformers (BERT) was employed in [5] with the purpose of performing word embedding and then integrated deep learning techniques with the purpose of predicting judicial cases judgment. This in turn resulted in the improvement of prediction accuracy significantly. A review of judgement classification and outcome based forecasting was investigated in [6]. However with focus on majority classes, performance results were not said to be satisfactory. To address this aspect, a novel label based attention mechanism and deep learning architecture to perform multi-label classification was proposed in [7]. With these mechanisms the error involved in classification was addressed, therefore paving way for both majority and minority classes extensively.

The applications of extracting information, machine learning and natural language processing have extended to various applications and hence possess considerable prospective for task automation in several areas. But, the difficulty of the tasks taking part in automation and the necessity of higher accuracy has restricted its implementation in the legal aspect.

Sentence Embedding and Capsule Networks were applied in [8] with the purpose of classifying legal documents into distinct classes. With the classified distinct classes, extractive summarizer employed it for generating document summary. According to the classified results were found to speed up the process involving legal document summarization. As a result the time involved in the summarization process was also said to be reduced considerably. However, it requires an enormous sample set to provide optimized results. In [9], a multi-task DL with transfer learning was applied that in turn overcame the issue of data scarcity in the legal domain. Though optimized results were obtained the method failed in considering the time consumed.

1.1 Contributory remarks

To address the issues mentioned in the above works, a Louvain Modularity and Ensemble Gradient Absolute Congruence (LM-EGAC) deep belief classifier for legal documents is proposed. This LM-EGAC method provides contributions as given below.

- To enhance the legal document classification precision rate, a LM-EGAC deep classifier is proposed taking into consideration optimized topic modeling and classification.
- A Highly Associative Unique Keyword and Louvain Modularity-based Optimized Topic Modeling are applied to LM-EGAC method for obtaining preprocessed legal documents based on highly associative IDF. Then, by applying the Louvain Modularity function, optimized topic words are extracted. This helps to improve the precision and recall involved during the classification process.
- Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier (EGBAC-DBC) is applied for legal document (i.e., judgment) classification. The Ensemble gradient boosting algorithm initially uses Absolute Congruence-based Deep Belief Classifier to classify the input legal judgment documents based on the extracted optimized topic word by applying Absolute Congruence function. The Ensemble gradient boosting algorithm then combines the classification performance via Binary Cross Entropy function therefore minimizing the overall time and error involved.
- The performance of the proposed method namely LM-EGAC is validated quantitatively and compared with other traditional methods. The results thus obtained demonstrate better performance in terms of precision, recall, time and error rate.

1.2 Organization of the work

Section 1 provides an introduction to the concept of legal document classification, its importance, application of ensemble learning, machine learning. In addition, this section examines the significance of classification of Indian judgment within the context of legal documents. Section 3 figures out the proposed method and its two-stage structures, contributing a broad overview of the method. Section 4 details on the experimental details of the proposed method. Section 5 introduces the results of the experiment and also detailed analysis of the outcomes by comparing traditional methods is included. Finally, Section 6 concludes the paper.

2. Related works

One of the most prevalent chores that legal personnel carry out each and every day and year is legal judgment. Also there arises a tradeoff between both the lawyers and clients in predicting lawsuits outcome anticipating to enhance their prospects of accomplishing results in one's favor.

An overview and review of deep learning techniques providing legal experts beneficial to both lawyers and clients were presented in [10]. However, multi-label classification for providing legal experts yet remains a major issue to be addressed. In [11], a fusion of stacked autoencoders and extreme learning machines were integrated via approximating the class score. This type of fusion mechanism resulted in the improvement of prediction capability to a greater extent. Yet another hierarchical multi-classification method was proposed in [12] using binomial neighborhood classifiers that in turn not only improved the prediction efficiency but also reduced the overall cost involved in performing the task.

A lot of development has been coerced in legal text processing over the past few decades. Amongst them text summarization is performed that refers to the process of obtaining the most significant segments of actual text and producing reasonable synopsis out of it. The objective for performing summarization of text in legal domain remains in assisting large amount of legal documents globally. However, such a task in recentyearsis said to be both labor and time consuming owing to the vast amount of data involved in it. Hence, legal document automatic summarization is considered to be one of the best solutions that in turn aid in assisting legal practitioners in a significant manner.

A multiclass perceptron algorithm employing multi labels was proposed in [13]. Also by applying the binary relevance mechanism resulted in the improvement of predictive performance to a greater extent. Yet another graph convolutional neural network along with pre-trained language models were applied in [14] to focus on the overhead and time aspects involved in classifying legal documents. However with multi-dimensionality of cases involved the prediction and training time were said to be compromised. To address this aspect, a classifier model using partial chain and highly correlative features using labeling method was proposed in [15]. Here, by learning the correlation between features assisted in the improvement of prediction accuracy with minimum convergence speed.

A weakly supervised approach was applied in [16] for documenting the Indian legal system. Also automatic labeling was also designed for document summarizing with improved precision. Yet another paper focusing on the optimization of the document summary in case of legal cases employing gravitational search algorithms was proposed in [17] and was proved to be better in terms of both precision and recall. A hybrid model was employed in [18] for documenting electronic medical records. Legal proceedings especially for women victims were focused on in [19] using machine learning. A state of the art method employing machine learning for legal documentation was investigated in [20].

With the effort to focus on the optimized topic modeling that can identify different cases involved in legal documents to make the readability easier, a method called Louvain Modularity and Ensemble Gradient Absolute Congruence (LM-EGAC) deep belief classifier for legal documents is proposed in this work. The elaborate description of the LM-EGAC method is provided in the following sections.

3. Louvain Modularity and Ensemble Gradient Absolute Congruence (LM-EGAC) deep belief classifier for legal documents

On the contrary to the courts in westernmost provinces, public documentations of Indian judiciary are not absolutely structured and hence found to be noisy. As a result till now there do not exist a large scale publicly accessible annotated datasets of Indian judiciary, therefore limiting the opportunity for legal analytics research. In this section an introduction of modeling a legal document employing Indian Supreme Court Judgements (https://www.kaggle.com/datasets/vangap/indian-supreme-court-judgments) using graph of words and extracting the pivotal words and sub graphs from it is presented. Let 'S' represent the sample instances and 'Res' represents the label space. Then, the task of multi-label classification to efficiently identify the case type in the given situation is to perform a mapping function as given below.

$$fun: S \rightarrow Res \rightarrow \{(S_i, Res_i) | \ 1 \leq i \leq C, \} \quad (1)$$

From the above equation (1), '$S_i \in S$', represents a sample instance of a situation and '$Res_i \subseteq Res$' represents the label sets associated with 'S_i'. Also for any undetectable sample instance '$S_i \in S$', multi-label classifier '$f(.)$' predict '$f(s) \subset Res$' as proper labels for 'S' respectively. With the above formulates the overall process of Louvain Modularity and Ensemble Gradient Absolute Congruence (LM-EGAC) deep belief classifiers for legal documents are designed in the following sections.

3.1 Highly Associative Unique Keyword and Louvain Modularity-based Optimized Topic Modeling

Natural Language Processing (NLP) provides an extensive scope of materials and methods for topic identification from legal documents. Traditional TF-IDF based model [1] employs the word distribution in the legal document for extracting topics. Nevertheless, these methods consider word neighborhoods and only extract exact word matches. We propose employing a variation of the Traditional TF-IDF based model called, Highly Associative Unique Keyword and Louvain Modularity-based Optimized Topic Modeling that by making the most of association between feature keywords and the volume of texts in which they emerge, while disregarding the discrepancy in feature keyword arrangement across classes identifies topics (i.e., top keywords) for obtaining predominant sections of the legal judgment. Fig 1 shows the structure of the Highly Associative Unique Keyword and Louvain Modularity-based Optimized Topic Modeling model.

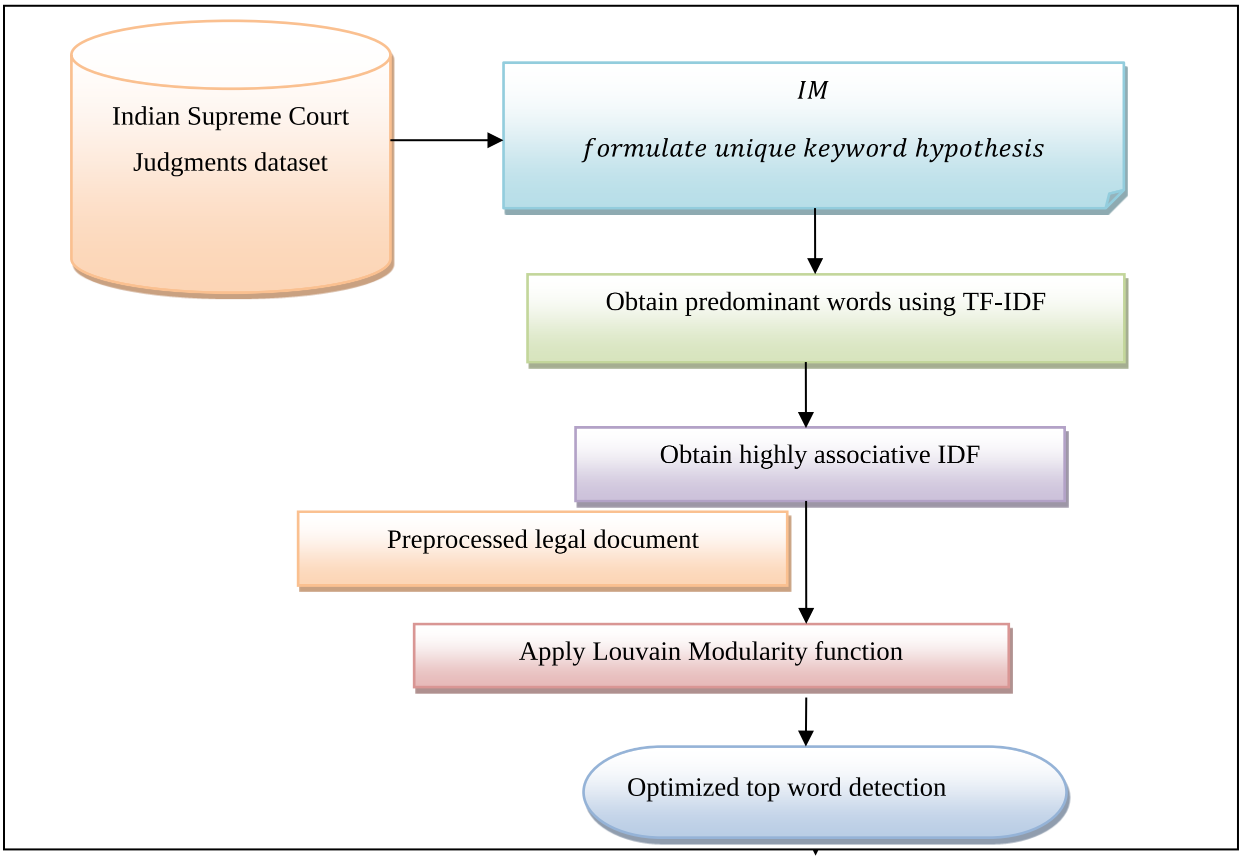

Fig 1. Structure of Highly Associative Unique Keyword and Louvain Modularity-based Optimized Topic Modeling

As illustrated in the above figure, let us consider the input matrix as given below with the raw Indian Supreme Court Judgments dataset as input.

$$IM = [S_1F_1S_1F_2\ \dots\ S_1F_nS_2F_1S_2F_2\ \dots\ S_2F_n\ \dots\ \dots\ \dots\ \dots\ S_mF_1S_mF_2\ \dots\ S_mF_n] \qquad (2)$$

From the above equation (2), the input matrix 'IM' is formulated based on the 'm' samples instances 'S_m' (i.e., legal judgments obtained as input) and 'n' features 'F_n' respectively. Let legal judgment '$LJ = \{LJ_1, LJ_2, \dots, LJ_m\}$' represent the set of '$m$' legal judgements obtained from the raw Indian Supreme Court Judgments dataset in the corpus. Moreover, '$m(LJ_i)$' denote the sentence set in legal document 'LJ_i' and let 'LJ_{ij}' denote the '$j - th$' sentence of the '$i - th$' legal judgment document. As the initial step in identifying unique keyword or legal document preprocessing are performed according to the presence/absence of nouns in the sample instance. While a person's name may have a lot of relevance in a news report, it just denotes a respondent in the corresponding case and hence does not have a major influence on any legal hypothesis in the judgment. As a result, it has to be discarded by means of unique keyword hypothesis set of the '$j - th$' sentence in the '$i - th$' legal document as given below.

$$UW(LJ_{ij}) = \{S \in LJ_{ij} | PoS(S) =' CN' and\ S \in PW(LJ_{ij})\} \qquad (3)$$

From the above equation (3), the unique word in the corresponding legal judgment of '$j - th$' sentence in the '$i - th$' legal document '$UW(LJ_{ij})$' is arrived at by modeling part of speech of the corresponding sample '$PoS(S)$' being subjected to predominant words '$PW(LJ_{ij})$' in the sentence 'LJ_{ij}' respectively. Here the common noun 'CN' Materializing in the sentences as hypothesis and establishing a hypothesis relationship graph employing hypothesis conjunctions are considered for evaluation. Nevertheless, only highly associative occurrences of nouns in the hypothesis graph are required to denote document segments. In our work, Highly Associative TF-IDF is designed to feature keywords in different cases (i.e., civil, criminal, sales tax and service) with distinct frequencies and interpretations in different cases.

Traditional TF-IDF measures weight for a keyword in a legal document by evaluating its local relevance employing term frequency (TF) within the legal document and global relevance employing inverse document frequency (IDF) for the entire legal document collection set. To validate the significance of the common noun, the traditional TF-IDF model is improvised by means of a highly associative mechanism for individual legal judgment taking into consideration each and every sentence in a legal judgment as an autonomous or independent legal document. Then, the predominant words '$PW(LJ_{ij})$' is determined as given below.

$$PW(LJ_{ij}) = S \in LJ_{ij} | TF(S, LJ_{ij}) * IDF(S, LJ_i) > \mu\left(TF(S, LJ_{ij}) * IDF(S, LJ_i)\right) \quad (4)$$

From the above equation (4), '$TF(S, LJ_{ij})$' represent the term frequency (TF) within the legal document of the corresponding sample 'S' in the legal judgment 'LJ_{ij}', whereas '$IDF(S, LJ_i)$' denotes the uniqueness of the corresponding sample 'S' across entire legal document collection set 'LJ_i'. Moreover, the keywords having '$TF - IDF$' greater than the 'μ' scores over the entire legal document collection set are considered as significant. Taking into consideration the traditional classification based on TF-IDF as a reference, by differentiating the meanings of feature keywords in different disciplines to enhance the confidence, highly associative Legal Document Unique Keyword Identification using IDF is mathematically stated as given below.

$$HA - IDF = - \log \log\left(1 - \frac{PW(LJ_{ij})(k_i)}{PW(LJ_{ij})(k_i) + PW(LJ_{ij})(O_i)}\right) \quad (5)$$

$$PR = \log \log\left(1 + \frac{PW(LJ_{ij})(k_i)}{PW(LJ_{ij})(O_i)}\right) \quad (6)$$

From the above equations (5) and (6), 'k_i', represents the probability of feature keyword 'i' being in class topic 'k' and 'O_i' represent the probability of feature keyword 'i' existing in other class topics 'O' in the legal document respectively. Here, other class topics in the legal document denote the classification results acquired after deleting topic class 'k' in the legal document.

Though the two feature keywords occur similar numbers of time in the legal document, however 'k_i' and 'O_i' are different owing to the diverging dispersal of definite feature keywords. Owing to this reason, the improved TF-IDF algorithm using Highly Associative Legal Document Unique Keyword Identification exceeds differentiation between distinct feature keyword arrangements. Detection of relevant keywords words in the legal judgment document is a predominant step and this is evaluated on the basis of closeness or vicinity of the unique keywords according to case (i.e., civil, criminal, service tax or service agreements). A case graph '$CG_i = (V_i, E_i)$' of a preprocessed legal document 'PR' is constructed with the aid of the unique keywords such that '$CG_i = (P_i, Q_i)$' of a legal document 'LJ_i' satisfies the following condition.

$$P_i = \cup_{j \in 1,..,m(LJ_i)} UW(LJ_{ij}) \quad (7)$$

$$Q_i = \{(S, Res) | Co - incidence(sen, Res) > 4\} \quad (8)$$

From the above equations (7) and (8) 'P_i' refers to the unique keywords across all sentences in the legal document and two keywords in the graph having an edge between unless their co-incidence is greater than 4. In other words, they occur simultaneously in no less than four sentences 'sen' respectively.

The frequency of co-incidence is employed in our work as the stability of association between two keywords. Here, in our work, less than 4 co-incidence of keywords are considered as scanty co-incidence and hence we considered them as less relevant and others are contemplated as high relevant unique keywords or topic words employed for further processing. The Louvain Modularity function for top word detection is a model to extract top words from large networks (i.e., large legal judgment sample). The top word to be optimized is referred to as modularity, ranging between '$-1/2$' and '1' that evaluates the density of links association inside top words compared to links between top words. This is performed using Louvain Modularity function as given below.

$$LM = \frac{1}{2m} \sum_{ij} \left[E_{ij} - \frac{l_i l_j}{2m}\right] \delta(c_i c_j) \quad (9)$$

From the above equation (9), 'E_{ij}' represents the weight of edge between legal documents 'i' and 'j', 'l_i' and 'l_j' refers to the sum of edge weights between legal documents 'i' and 'j', 'c_i' and 'c_j' forms the number of cases with respect to legal documents and finally 'δ' denoting the kronecker delta function (i.e., '$\delta(c_i c_j) = 1, if\, c_i = c_j, 0\ otherwise$'. Finally, based on the above equation (9), the modularity of a specific case (i.e., civil, criminal, sales tax, service) is then mathematically stated as given below.

$$LM_c = \frac{\Sigma_{in}}{2m} - \left(\frac{\Sigma_{tot}}{2m}\right)^2 \quad (10)$$

From the above equation (10), 'Σ_{in}' refers to the sum of edge weights between legal documents within the case and 'Σ_{tot}' represents the sum of all edge weights for legal documents within the case respectively. With this, the optimized topic words are returned in a precise manner. The pseudo code representation of the Highly Associative Unique Keyword and Louvain Modularity-based Optimized Topic Modeling model is given below.

Input: Dataset 'DS', Samples '$S = \{S_1, S_2, \dots, S_m\}$', Features '$F = \{F_1, F_2, \dots, F_n\}$'

Output: Precise and accurate optimized topic modeling

Step 1: **Initialize** 'm', 'n'

Step 2: **Begin**

Step 3: **For** each 'DS' with 'S' and 'F'

//Preprocessing

Step 4: Formulate a mapping function $fun: S \rightarrow Res \rightarrow \{(S_i, Res_i) | 1 \leq i \leq C, \}$

Step 5: Formulate input matrix $IM = \begin{bmatrix} S_1F_1 & S_1F_2 & \dots & S_1F_n \\ S_2F_1 & S_2F_2 & \dots & S_2F_n \\ \dots & \dots & \dots & \dots \\ S_mF_1 & S_mF_2 & \dots & S_mF_n \end{bmatrix}$

Step 6: Evaluate unique keyword hypothesis $UW(LJ_{ij}) = \{S \in LJ_{ij} | PoS(S) =' CN' and\ S \in PW(LJ_{ij})\}$, $PW(LJ_{ij}) = S \in LJ_{ij} | TF(S, LJ_{ij}) * IDF(S, LJ_i) > \mu\left(TF(S, LJ_{ij}) * IDF(S, LJ_i)\right)$

Step 7: Evaluate highly associative Legal Document Unique Keyword Identification using IDF as given in $HA - IDF = -\log\left(1 - \frac{PW(LJ_{ij})(k_i)}{PW(LJ_{ij})(k_i) + PW(LJ_{ij})(O_i)}\right)$

Step 8: **Return** preprocessed results $PR = \log\left(1 + \frac{PW(LJ_{ij})(k_i)}{PW(LJ_{ij})(O_i)}\right)$

//Feature extraction

Step 9: **For** each 'PR'

Step 10: Construct case graph $P_i = \cup_{j \in 1,..,m(LJ_i)} UW(LJ_{ij})$, $Q_i = \{(S, Res) | Co - incidence(sen, Res) > 4\}$

Step 11: Apply Louvain Modularity function $LM = \frac{1}{2m} \sum_{ij} \left[E_{ij} - \frac{l_i l_j}{2m}\right] \delta(c_i c_j)$

Step 12: Evaluate modularity of a specific case $LM_c = \frac{\Sigma_{in}}{2m} - \left(\frac{\Sigma_{tot}}{2m}\right)^2$

Step 13: **Return** optimized extracted topic words 'LM_c'

Step 14: **End for**

Step 15: **End for**

Step 16: **End**

Algorithm 1 Highly Associative Unique Keyword and Louvain Modularity-based Optimized Topic Modeling model

As given in the above algorithm, an optimized topic modeling extracting optimized topic keywords is designed. Here, two processes are involved, namely preprocessing and feature extraction. With the raw dataset acquired as input, first, the features and samples are subjected to preprocessing to obtain a highly associative unique keyword. For obtaining a highly associative unique keyword, the traditional TD-IDF is modified according to the frequency of co-incidence with respect to the stability of association between keywords. Next, optimized topic modeling or optimized topic words are extracted using the Louvain Modularity function. The function with no prior assumptions required to construct a case graph retrieves precise features or topics in an accurate fashion.

3.2 Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier

The proposed method uses Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier (EGBAC-DBC) for classifying the legal documents with optimized extracted topic words. In our work, a machine learning framework called ensemble learning is used that trains multiple results called weak learners with which the results are integrated to obtain optimal solutions. Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier employed in our work is a type of supervised ensemble classification that combines learners with poor performance. The term gradient boost refers to an ensemble technique that aggregates classification results with the objective of providing accurate and reliable results in a minimum time and overhead. Fig 2 shows the structure of the Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier (EGBAC-DBC) model.

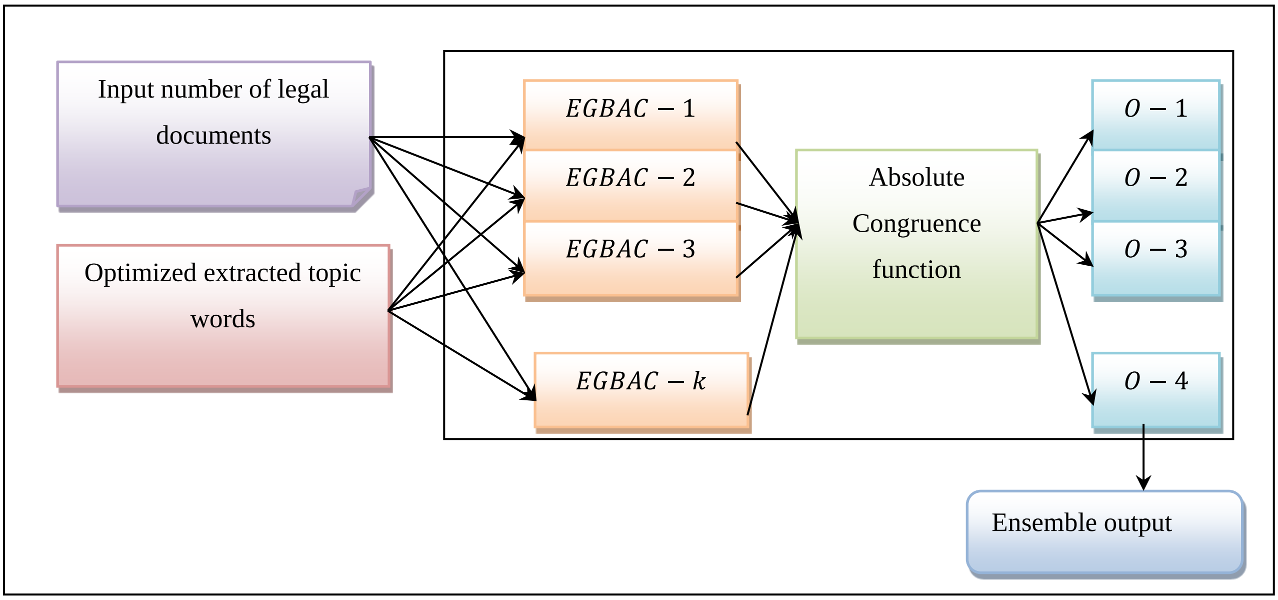

Fig 2. Structure of Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier (EGBAC-DBC) model

Fig 2 given above illustrates the process of Structure of Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier (EGBAC-DBC) model to generate the final results of classification. The ensemble classifier in addition to the training sets as 'S_i, Res_i' where 'S_i' stands for the sample legal documents and 'Res_i' takes into account the optimized extracted topic words to obtain the results of the ensemble classification for the inputs that were provided. The ensemble method creates 'k' number of Ensemble Gradient Boost Absolute Congruence to classify the given input legal documents along with optimized extracted topic words. The proposed Absolute Congruence-based Deep Belief Classifier architecture comprises three units namely input unit, output unit and two hidden units. The structure of Absolute Congruence-based Deep Belief Classifier is shown in fig 3.

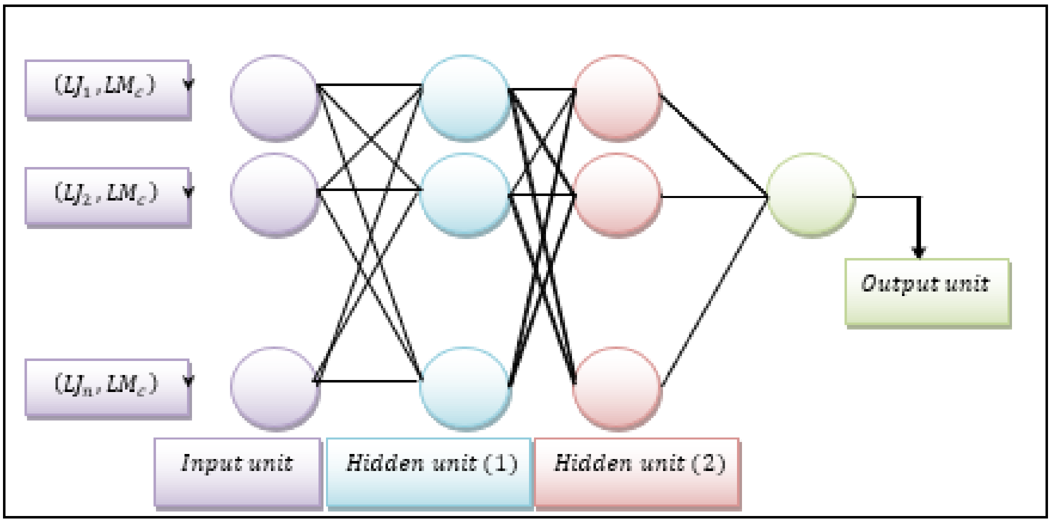

Fig 3. Structure of Absolute Congruence-based Deep Belief Classifier

As illustrated in the above figure, with one input unit, two hidden units and one output unit, initially in addition to legal documents optimized extracted topic words are fed as input to the initial layer. With this optimized extracted topic words only, the corresponding legal documents via processing in two hidden units identify the type of case in the given situation in a feed forward manner to model the entire network. Following which the first hidden layer transforms the given input wherein Absolute Congruence function is applied to identify highly correlated results towards efficient multi label classification.

One problem with the Spearman's rank correlation coefficient [2] is that it is sensitive to changes in sign due to the reason that when only 'LJ_i' or 'c_j' change in sign then the sign of the product changes. But '$\varphi(LJ_i, c_j)$' overestimates congruence when the signs of the variable pairs '(LJ_i, c_j)' is predominantly the same and also the similarity is underestimated if the signs are predominantly different, therefore consuming a significant amount of time and overhead while performing ensemble function and hence the overall classification results. To address this aspect, Absolute Congruence function is applied in our work in the second hidden unit and is formulated as given below.

$$AC_o = \varphi(LJ_i, c_j) = \frac{\sum|LJ_i c_j|}{\sqrt{\sum LJ_i^2 \sum c_j^2}} \tag{11}$$

From the above equation (11), the value of '$\varphi(LJ_i, c_j : LM_c)$' ranges from '$-1$' to '$+1$' indicating Absolute Congruence value, 'LJ_i' refers to legal documents and 'c_j' refers to a specific case. The congruence value ranging from '-1' (i.e., negative congruence) to '$+1$' (i.e., positive congruence) is then passed to the second hidden layer. The congruence correlation coefficient value ranges from -1 to +1. According to the congruence results, the legal documents suitably classified into their respective cases with minimum time. On the other hand, the ensemble model boosts the classification results by joining several deep learning classifiers. This is formulated using the Binary Cross Entropy function as given below.

$$\mathrm{W_i} = \min_{\Theta} -\frac{1}{n}\sum_{c=1}^{n}\sum_{o=1}^{m}\left[LM_c \log(\sigma(AC_o)) + (1 - LM_c)\log(1 - \sigma(AC_o))\right] \tag{12}$$

From the above equation (12) results, it is identified that the Binary Cross Entropy based loss function is more suitable for multi-label classification therefore reducing considerable time in generating case types in an accurate manner. Finally, the error rate is mathematically formulated as given below.

$$WC_E = \{exp\ exp(W_i) - act(W_i)\} \tag{13}$$

From the above equation (13), 'WC_E' refers to the weak classifier error rate measured using the expected results '$exp\ exp(W_i)$' and the actual results '$act(W_i)$' respectively. The weight of each deep classifier is reassigned in such a manner so as to reduce the error, therefore obtaining accurate classified results for each legal document. The algorithmic process of Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier is as given below.

Input: 'DS', '$S = \{S_1, S_2, \ldots, S_m\}$', '$F = \{F_1, F_2, \ldots, F_n\}$'
Output: computationally efficient and precise classifier results
Step 1: **Initialize** 'm', 'n', 'LM_c' Step 2: **Begin** Step 3: **For** each 'DS' with 'S', 'F' and 'LM_c' Step 4: Construct 'k' deep learning classifier set Step 5: Obtain 'S' (i.e., legal judgment documents) and 'LM_c' in the input layer Step 6: Apply Absolute Congruence function $AC_o = \varphi(LJ_i, c_j) = \frac{\sum\lvert LJ_i, c_j\rvert}{\sqrt{\sum LJ_i^2 \sum c_j^2}}$ Step 7: Formulate weight function $\mathrm{W_i} = \min_{\Theta} -\frac{1}{n}\sum_{c=1}^{n}\sum_{o=1}^{m}\left[LM_c \log\left(\sigma(AC_o)\right) + (1 - LM_c)\log\left(1 - \sigma(AC_o)\right)\right]$ Step 8: **If** 'S' is closer to 'LM_c' Step 9: Classify legal judgment documents into a particular case Step 10: Combine all weak classifier results Step 11: Update weight $WC_E = \{\exp(\mathrm{W_i}) - act(\mathrm{W_i})\}$ Step 12: **End if** Step 13: **Return** classified legal judgment documents into particular case Step 14: **End for** Step 15: **End**

Algorithm 2 Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier

As given in the above Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier algorithm the weak classifiers are ensembles to build a strong classifier. To start with, the initialization of weights is done with arbitrary values for each deep learning classifier. Accordingly the error rates are measured and weights are updated for various iterations. With this the overall ensemble process validates the best classifier with minimum error. Owing to this also the time consumed in the overall classification process is minimized.

4. Experimental settings

The proposed Louvain Modularity and Ensemble Gradient Absolute Congruence (LM-EGAC) deep belief classifier for legal documents and existing works Latent Dirichlet Allocation-based natural language processing [1] and Multi-label Classification using ML [2] are tested. The work is implemented using Python, a high-level general-purpose programming language. The dataset used for the work is obtained from Kaggle, called Indian Supreme Court Judgment Dataset (https://www.kaggle.com/datasets/vangap/indian-supreme-court-judgments). A collection of 45000 legal case documents with judgments extracted from the government website API are used for the work. The case documents starting from 1950 and as of 30-11-2022 are included in the dataset wherein classification of four distinct types of cases, namely, civil, criminal, sales tax and service are performed for analysis.

5. Discussion

The experimental evaluation of the proposed Louvain Modularity and Ensemble Gradient Absolute Congruence (LM-EGAC) deep belief classifier method and existing works Latent Dirichlet Allocation-based natural language processing [1] and Multi-label Classification using ML [2] are compared with the aid of metrics like, precision, recall, training time and error rate. The performance of the above three methods is visualized with the aid of tables and graphs. The precision rate is measured as given below.

$$Pre = \frac{TP}{TP+FP} * 100 \quad (14)$$

From the above equation (14), the precision rate 'Pre', is obtained by taking the true positive rate 'TP' (i.e., truly classifying the type of case as it is) and false positive rate 'FP' (i.e., incorrectly predicting the positive class).

Table 1 Comparative analysis of precision using LM-EGAC, Latent Dirichlet Allocation-based natural language processing [1] and Multi-label Classification using ML [2]

Number of legal documents	**Precision (%)**		
	LM-EGAC	Latent Dirichlet Allocation-based natural language processing	Multi-label Classification using ML
4000	93.75	90	88.75
8000	94.15	88.15	77.55
12000	96	89	82
16000	98.35	89.35	82.35
20000	94.25	87.45	80.25
24000	95	88.35	81
28000	95.35	88.85	81.05
32000	96.35	89	81.35
36000	97	89.15	82
40000	97.55	90	82.35

The performance of the proposed LM-EGAC method in terms of precision and its comparative analysis with other two existing methods [1] [2] is given in Table 1 for which the results are arrived at when substituting the values in equation (14). The convergence graph is neither increasing nor decreasing with the increase in legal documents provided as input. From the observed results it is inferred that the precision rate when substituted using equation (14) is found to be higher using LM-EGAC method upon comparison with the existing methods [1] and [2]. As provided in table 1, for a sample size of 4000 legal documents, the results of precision using LM-EGAC method was found to be 93.75%. On the other hand, the precision of Latent Dirichlet Allocation-based natural language processing [1] and Multi-label Classification using ML [2] was found to be 90% and 88.75% respectively. A simulation of ten iterations were performed with 40000 different legal documents and the validation performed shows that the precision of LM-EGAC method is improved by 8% and 18% LM-EGAC method compared to [1] and [2]. The reason behind the improvement was the acquisition of Highly Associative Unique Keywords as preprocessed results from the raw dataset. With this the discrepancy arising in feature keywords across classes are discarded from further processing. This in turn improves the true positive rate, therefore improving the precision rate using LM-EGAC method by 8% compared to [1] and 14% compared to [2].

Second recall rate is mathematically formulated as given below.

$$Rec = \frac{TP}{TP+FN} * 100 \quad (15)$$

From the above equation (15), recall 'Rec' is measured using the true positive rate 'TP' and the false negative rate 'FN' (i.e., incorrectly predicting the negative class) respectively.

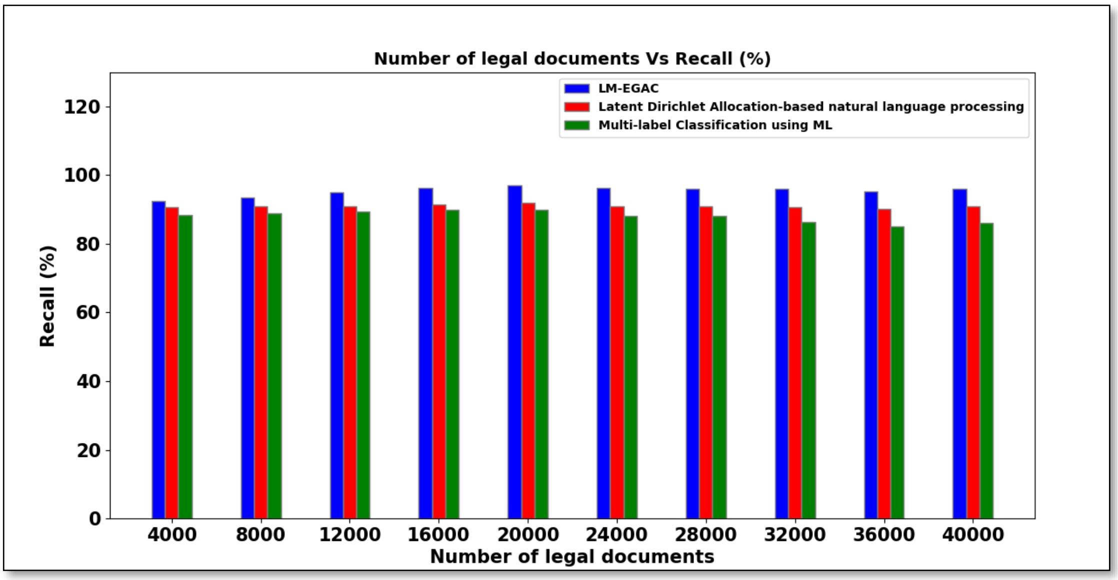

Fig 4. Recall versus legal documents

The performance of the proposed LM-EGAC method in terms of recall and its comparison with other two methods Latent Dirichlet Allocation-based natural language processing [1] and Multi-label Classification using ML [2] is shown in Fig 4. To ensure fair comparison similar sets of legal documents are provided as input in all the three methods, following which ten iterations were performed to analyze the performance of three methods to observe the recall rate. From the above graphical and tabular result it is inferred that the rate of recall of LM-EGAC method is better than the other two existing methods [1] and [2]. Let us consider a scenario of 4000 legal documents, the recall rate obtained using the LM-EGAC method was 92.36%, 90.56% and 88.39% were observed using [1] and [2]. This is due to the application of the Ensemble Gradient Absolute Congruence boosting model for classifying the legal documents that in turn increases the true positive rate by decreasing the false negative rate simultaneously. Also by applying the Louvain Modularity-based Optimized Topic Modeling model optimized topic words are extracted that in turn aids in decreasing the false negative. With this the overall recall rate using LM-EGAC method is said to be improved by 5% compared to [1] and 8% compared to [2].

Third training time involved in legal document classification according to the type of cases is analyzed. Training time measures the time consumed in classifying the legal documents according to their corresponding types of cases and is measured as given below.

$$TT = \sum_{i=1}^{m} S_i * Time\ (Class[LJ]) \quad (16)$$

From the above equation (16), training time 'TT' is analyzed by taking into consideration the sample legal judgment documents 'S_i' and the time consumed in classification '$Time\ (Class[LJ])$'. It is measured in terms of milliseconds (ms). Table 2 given below lists the training time involved both with and without preprocessing.

Table 2 Comparative analysis of training time (with and without preprocessing)

Number of legal documents	**Training time –without preprocessing (ms)**			**Training time – with preprocessing (ms)**		
	LM-EGAC	Latent Dirichlet Allocation-based natural language processing	Multi-label Classification using ML	LM-EGAC	Latent Dirichlet Allocation-based natural language processing	Multi-label Classification using ML
4000	220	288	340	160	220	272
8000	255	300	365	215	245	315

12000	310	335	435	245	335	355
16000	335	395	485	280	350	385
20000	385	450	550	315	415	435
24000	425	515	585	355	455	500
28000	460	535	595	370	480	525
32000	485	585	610	400	500	545
36000	535	635	685	425	535	590
40000	585	680	735	450	600	635

Table 2 given above shows the training time using three methods namely the LM-EGAC method and existing Latent Dirichlet Allocation-based natural language processing [1] and Multi-label Classification using ML [2]. The training time here refers to the time consumed for classifying the given preprocessed legal document into the type of cases (i.e., civil, criminal, sales tax, service) based on the optimized topic modeling. Also to ensure fair comparison between the proposed and existing methods, ten iterations were performed using a similar set of input legal judgment documents with which the graph was plotted. Similar analysis was performed both with preprocessing and without preprocessing and results were validated separately. Using the LM-EGAC method, when 4000 legal documents along with optimized topic modeling were provided as input, the training time involved was observed to be 160ms, whereas 220ms and 272ms using the existing methods (after preprocessing). However, without preprocessing the training time involved using LM-EGAC method was observed to be 220ms, 288ms and 340ms using [1] and [2]. From these results two things are inferred, first, the training time with preprocessing was found to be better than without preprocessing, therefore corroborating the objective considered in the LM-EGAC method. In a similar manner, the training time both with and without preprocessing using LM-EGAC method was comparatively lower than [1] and [2]. The reason was due to the application of Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier. By applying this algorithm Absolute Congruence function was in the first hidden layer that in turn efficiently differentiated between the positive and negative congruence function. As a result, the training time involved in the overall classification process using LM-EGAC method was found to be reduced by 15% [1] and 27% [2] for without preprocessing and 22% [2] and 30% [2] for with preprocessing.

Finally, classification error involved is mathematically represented as given below. While performing a certain amount of error is said to occur and hence is said to be inevitable.

$$CE = \sum_{i=1}^{m} \frac{S_{WC}}{S_i} * 100 \tag{17}$$

From the above equation (17), the classification error 'CE' is measured based on the sample legal judgment documents 'S_i' and the sample legal judgment documents wrongly classified 'S_{WC}'. It is measured in terms of percentage (%). Lower the classification error, effective the method is said to be. Fig 4 lists the classification error using the three methods.

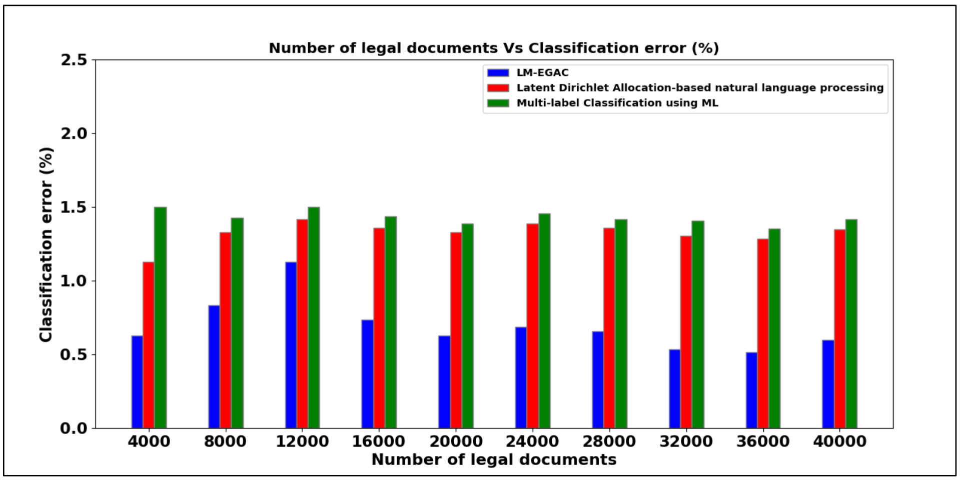

Fig 5. Classification errors versus legal documents

Fig 5 given above illustrates the classification error involved in the overall process of identifying the type of case with the legal documents and optimized topic modeling provided as input. From the above figure neither increasing nor decreasing trend is observed using all the three methods for samples ranging between 4000 and 40000. Also with an overall sample size of 40000 the error rate was 0.6% using LM-EGAC and 1.345% and 1.415% using [1] and [2]. Also the highest classification error of 0.735% was observed when the samples were 16000 using LM-EGAC, 1.355% and 1.435% when applied with [1] and [2]. From this results the classification error rate using LM-EGAC method was lesser than [1] and [2]. The application of Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier algorithm aided in attaining this objective. By applying Binary Cross Entropy function was applied in the second hidden layer for updating the weight. This assisted in minimizing the samples from wrongly classifying using LM-EGAC method. As a result the overall classification error using LM-EGAC method was found to be less by 48% compared to [1] and 52% compared to [2].

6. Conclusion

In this work a method called Louvain Modularity and Ensemble Gradient Absolute Congruence (LM-EGAC) is designed that assists in identifying the type of case provided with legal documents. The LM-EGAC initially performs the optimized topic modeling using the Highly Associative Unique Keyword and Louvain Modularity-based Optimized Topic Modeling for extracting optimized top words. Next, Ensemble Gradient Boost Absolute Congruence-based Deep Belief Classifier is applied for document classification according to the type of case with extracted optimized top words. The Ensemble boosting algorithm uses the Absolute Congruence-based Deep Belief Classifier to classify the legal documents. Strong classifier results are hence obtained with comparatively less error. To analyze the performance of LM-EGAC method significant performance parameters like, precision, recall, training time and error rate are measured. According to the findings, the LM-EGAC method achieves higher levels of precision with minimum time and error than the other traditional methods.

References

[1] Isha Gupta, Indranath Chatterjee, Neha Gupta, "A two-staged NLP-based framework for assessing the sentiments on Indian supreme court judgments", International Journal of Information Technology, Springer, Apr 2023 [Latent Dirichlet Allocation-based natural language processing]

[2] Francisco de Arriba-Perez, Silvia Garcia-Mendez, Francisco J. Gonzalez-Castano, Jaime Gonzalez-Gonzalez, "Explainable machine learning multi-label classification of Spanish legal judgements", Journal of King Saud University – Computer and Information Sciences, Elsevier, Oct 2022

[3] Naseer Ahmed Sajid, Munir Ahmad, Atta-ur Rahman, Gohar Zaman, Mohammed Salih Ahmed, Nehad Ibrahim, Mohammed Imran B. Ahmed, Gomathi Krishnasamy, Reem Alzaher, Mariam Alkharraa, Dania AlKhulaifi, Maryam AlQahtani, Asiya A. Salam, Linah Saraireh, Mohammed Gollapalli and Rashad Ahmed, "A Novel Metadata Based Multi-Label Document Classification Technique", Computer Systems Science & Engineering, Techscience, Nov 2022

[4] Arousha Haghighian Roudsari, Jafar Afshar, Wookey Lee, Suan Lee, "PatentNet: multi-label classifcation of patent documents using deep learning based language understanding", Scientometrics, Springer, Feb 2022

[5] Yongjun Wang, Jing Gao, Junjie Chen, "Deep Learning Algorithm for Judicial Judgment Prediction Based on BERT", IEEE Xplore, Dec 2020

[6] Masha Medvedeva, Martijn Wieling, Michel Vols, "Rethinking the feld of automatic prediction of court decisions", Artifcial Intelligence and Law, Springer, Jan 2022

[7] Dezhao Song, Andrew Vold, Kanika Madan, Frank Schilder, "Multi-label legal document classification: A deep learning-based approach with label-attention and domain-specific pre-training", Multi-label legal document classification: A deep learning-based approach with label-attention and domain-specific pre-training, Elsevier, Jan 2021

[8] Harshith R. Acharya, Aditya D. Bhat, K. Avinash and Ramamoorthy Srinath PES University, Bengaluru, India, "LegoNet - classification and extractive summarization of Indian legal judgments with Capsule Networks and Sentence Embeddings", Journal of Intelligent & Fuzzy System, Springer, Oct 2022

[9] Ahmed, Christoph, Ingo, and Florian, "Multi-Task Deep Learning for Legal Document Translation, Summarization and Multi-Label Classification", ACM Digital Library, Sep 2018

[10] Elias Jacob, Marco Bruno Miranda Clementino, “Using deep learning to predict outcomes of legal appeals better than human experts: A study with data from Brazilian federal courts”, PLOS ONE | https://doi.org/10.1371/journal.pone.0272287 July 28, 2022

[11] Anwesha Law, Ashish Ghosh, “Multi-label classification using a cascade of stacked autoencoder and extreme learning machines”, Neurocomputing, Elsevier, May 2019

[12] Miguel Romero, Jorge Finke and Camilo Rocha, “A top-down supervised learning approach to hierarchical multi-label classifcation in networks”, Applied Network Science, Springer, Oct 2022

[13] Eneldo Loza Mencıa and Johannes F¨urnkranz, “Efficient Multilabel Classification Algorithms for Large-Scale Problems in the Legal Domain”, Springer, Feb 2011

[14] Bin Liu, “GCN-BERT and Memory Network Based Multi-Label Classification for Event Text of the Chinese Government Hotline”, IEEE Access, Oct 2022

[15] Zhenwu Wang, Tielin Wang, Benting Wan and Mengjie Han, “Partial Classifier Chains with Feature Selection by Exploiting Label Correlation in Multi-Label Classification”, Entropy, Oct 2022

[16] Vedant Parikh, Vidit Mathur, Parth Mehta, Namita Mittal , Prasenjit Majumder, “LawSum: A weakly supervised approach for Indian Legal Document Summarization”, DeepAI, Oct 2021

[17] Ambedkar Kanapala, Srikanth Jannu, Rajendra Pamula, “Summarization of legal judgments using gravitational search algorithm”, Neural Computing and Applications, Springer, Mar 2019

[18] Tingzhong Wang, Yongxin Zhang, Yifan Zhang, Hao Lu, Bo Yu, Shoubo Peng, Youzhong Ma, and Deguang Li, “A Hybrid Model Based on Deep Convolutional Network for Medical Named Entity Recognition”, A Hybrid Model Based on Deep Convolutional Network for Medical Named Entity Recognitio, Hindawi, May 2023

[19] Elena Escobar-Linero, Marı´a Garcı´a-Jime´nez, Marı´a Eva Trigo-Sa´nchez, Marı´a Jesu´s Cala-Carrillo, Jose´ Luis Sevillano, Manuel Domı´nguez-Morales, “Using machine learning-based systems to help predict disengagement from the legal proceedings by women victims of intimate partner violence in Spain”, PLOS ONE, Jun 2023

[20] Varun Dogra, Sahil Verma, Kavita, Pushpita Chatterjee, Jana Shafi, Jaeyoung Choi, and Muhammad Fazal Ijaz, “A Complete Process of Text Classification System Using State-of-the-Art NLP Models”, Computational Intelligence and Neuroscience, Hindawi, Jun 2022

[21] Y. Hole, S. Hole, L. P. Leonardo Cavaliere, B. Nair, M. Hasyim and H. B. Bapat, (2023) "Blockchain Usages in Hospitality Management," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023, pp. 2798-2801, doi: 10.1109/ICACITE57410.2023.10183291.

[22] Y. Hole, S. Hole, A. A. Ayub Ahmed, E. Efendi, I. Ibrahim and M. Hasyim, (2023) "Internet of Things Issues and Challenges," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2023, pp. 1370-1373, doi: 10.1109/ICACITE57410.2023.10183221.